INFANT MORTALITY STUDIES OF THE CHILDREN'S BUREAU.*

By Robert M. Woodbury.

Infant mortality studies have been made by the Children's Bureau in eight cities,† characterized by various and industrial and social conditions. The reports for Johnstown, Pa., and Manchester, N. H., have already been published. Besides these, field studies have been made in Saginaw, Mich., Waterbury, Conn., Brockton and New Bedford, Mass., Akron, Ohio, and Baltimore, Md. Of these cities Baltimore is in many respects the most interesting since it contains a large negro population; the data here are for a large number of births and deaths and offer therefore a basis for valid conclusions on many interesting problems.

The object of these studies is to show the causes of infant mortality and the relation between the mortality rates and social, economic, and physical‡ factors. For this purpose as well as for the purpose of comparing rates under different industrial conditions, infant mortality rates must be available that are substantially accurate. Records of births and deaths of the Census Bureau are indeed available for calculating rates for cities in the birth registration area, but the margin of error varies from city to city, and probably from class to class in the same city; further, many points of information that were deemed important in a study of causes are not included in the registration records. The method adopted, therefore, was to supplement the registration records by inquiries covering type of feeding, earnings of father, color and nativity of mother, employment of mother, housing conditions, age of mother at birth, order of birth, maternal history, etc. With the results obtained it became possible for the first time in this country to study the effect of these various factors upon the mortality under one year of age.§

* Paper read at the seventy-ninth Annual Meeting of the American Statistical Association.

† Exclusive of Montclair, N. J., already published. In Montclair, the schedules were furnished by the Bureau, and the data tabulated, but the information was not gathered by the Bureau's agents.

‡ Sex of infant, single and plural birth, prematurity, age of mother, order of birth.

§ With the possible exception of the study made by the Bureau of Labor on Infant Mortality and its Relation to the Employment of Mothers, Vol. xiii of Report on Condition of Woman and Child Wage Earners in the United States, 61st Cong., 2nd Sess., Sen. Doc. No. 645. The Massachusetts statistics did not permit of anything more than unsatisfactory general correlations; the Fall River investigation was based on a study of deaths only.

An infant mortality rate expresses the probability of a live born infant dying before the first birthday per 1,000 live born infants. The usual approximate method of calculation is to compare all deaths in the city in a given year to all births in the city in the same interval—a method that is open to the objection that the deaths in the city may include some deaths of infants born elsewhere and some of infants born in the preceding year. When deaths and births are subclassified by race, nationality, economic status of father, etc., the effect of these inaccuracies becomes difficult to estimate. Accordingly the method followed by the Bureau in all of its studies has been to follow each infant born in a given selected year from birth till death or till the first birthday was reached. In this way, in theory, for each nationality or economic or other subclass a probability of dying in the first year of life is secured, each rate accurate and comparable with the other rates obtained.

In practice, however, difficulties arose. How were all births in the city in the year selected to be discovered? The obvious method was to search the birth records, and the facility afforded by these records was an important factor in the choice of cities to be investigated. But these records are faulty; in some cities a considerable percentage of births is not registered, in others only a few are omitted. Some evidence of faulty registration is given in the death records themselves; a death record may be at hand of an infant (born in the city) for whom no birth record is available. Records are probably more complete in a state like Massachusetts where certification is carefully checked every year than in a state where both the idea of registration is new and the officials have not learned the administrative methods necessary to secure complete registration. In the absence of a careful canvass it may be difficult to determine how complete the records of births are. Records of deaths are likewise often incomplete, though in general laws requiring death registration are more widely adopted and better enforced than those for birth registration.* If exactly the same percentage of deaths as of births is omitted, in each

*In Saginaw 3 deaths among 118 were found to have escaped registration; 149 births were added to 1,113 that had been registered.

class studied, the mortality rates will still be correct; but in most cases these percentages will not be identical, and usually the percentage of omissions of births will exceed that for deaths. In any case it is important to discover in what direction the errors in the mortality rates lie, and what procedure may best reduce or eliminate these errors or safeguard the conclusions.

In other words, the varying incompleteness of registration, which affects not only the general rate for the city but also the rate for each subclass and possibly also the comparability of these rates, is a serious difficulty. The question is put whether, by selecting a certain group for study, it would be possible to secure complete records for that group that would be somewhat more reliable than records for the city. The groups where the rates are most open to question and the facts most difficult to obtain are the births to migratory and non-resident mothers and the illegitimate births.

The migratory group where the family has moved away from the city is especially difficult to follow. Thorough and complete investigation would multiply the expense of inquiry, if indeed it were possible in all cases to find the mother to secure the information. It is certainly admissible to treat this migratory group separately; separate treatment seems especially advisable in view of the fact that the deaths in the group of infants born in the city and removed is not a complete tally because only the early deaths are certain to be registered. It would be interesting to know whether the fact that the family moved or that the family is a migratory one has any influence on mortality rates, but as yet the information secured has seemed so unsatisfactory that the results for this group have not been published.

One consideration upon which this policy of excluding migratory families was based was that the rates were to characterize the city; that housing conditions inseparable from geographical location, climate, etc., and especially local industrial, economic and social conditions materially affected mortality. Strictly speaking, the numerator of the fraction expressing the mortality rate should then include all deaths of infants in the locality, and the denominator the total number of years lived by all infants in the city. But in case of infants born else-

where, the mortality and health might depend upon or be influenced by conditions elsewhere, and the infants born to non-resident mothers might have either better or poorer chances of life due to conditions not peculiar to the locality. The inclusion of such cases among the deaths and the life time corresponding might therefore slightly influence the rates. In one class of cases in particular this influence would be marked; special hospital maternity cases, where infant mortality is considerably more than average, may include many non-resident mothers. Accordingly, to make the result as characteristic of the city as possible, all births to non-resident mothers were excluded. But in the main, the mortality rates are determined more by the type of feeding, economic class, etc., and the proportion of each class in the city, than by geographic and purely local climatic conditions.

A further reason for the exclusion or separate treatment of migratory families is the difficulty of securing an adequate statement of the years lived in the city. There are no data of the average age of infants moving away nor, for that matter, of the average age of infants moving in from the outside. If this information were available it would be possible to make monthly rates of mortality and to compare the monthly rates for migratory families with the rates where the family remained in the city.

A corollary of this policy of exclusion is that deaths in the city of infants born to mothers who later moved away prior to the visit of the agent have to be excluded. If not, the policy of exclusion would subtract only from the life time lived in the city without any deaths to correspond. A special class of cases is where the mother died. These cases are included if the infant or the family was living in the city at the time of the visit of the agent and excluded if the family had moved away.

The illegitimate births present certain special problems. There are difficulties in securing accurate information, for the families are frequently "non-resident" or have moved away. Nor are the data secured as trustworthy as for legitimate births. Furthermore, the conditions surrounding these infants are not those of the normal family, and these births should therefore be given separate treatment. In some cities no at-

tempt was made to secure schedules for the illegitimate births. In Baltimore the subject will be treated in a separate chapter.

The data presented then are reasonably complete and accurate for the legitimate births in the selected year to resident mothers living in the city at the time of the investigation.

For the rest, with migrant and non-resident and illegitimate groups excluded, three general methods of selection of the group are available. One is to take the records as they are, treating each death record as a birth record if the birth actually occurred in the city. The second is to take records of births as the starting point and accept records of deaths only when the birth record is available. The third is to supplement the records of births and deaths by a thorough canvass.

The three methods will give exactly the same results if the records are complete. If the records are not complete, the choice of method may be affected by the character and extent of the omissions. If death registration is complete but birth registration faulty, the first method will obviously give too high rates; if birth registration is nearly complete among the native but faulty among the foreign born, the method will exaggerate the differences between the infant mortality rates for the nativity classes.

The second method is an attempt to avoid the bias of the first method by the selection of births among whom the deaths are complete. The rates for each class of selected births are formed by comparing the number of infants that died (birth registered) with the sum of the infants that died plus those that survived the first birthday. These rates are correct in the sense that it is ascertained for each legitimate infant (of the resident group) whose birth was registered whether the infant survived or did not survive at the first birthday. Whether the rates are typical of the group is another and quite as vital a question; are enough infants that died associated with the infants that lived in each group to make a typical rate? Births in the more well-to-do families are more likely to be registered than in poorer families; the mortality rate for the group of infants selected on the basis of registration of birth tends, therefore, to be somewhat lower than the rate for all births. At first thought, it might appear that, in spite of this tendency

away from a rate typical of the entire group, the rates for each subgroup would be typical and allow of safe comparisons. But even here, the same tendency may show itself, the births in the more intelligent and better situated subgroups being more likely to be registered and therefore with fewer deaths among them to be associated with the registered births. Other reasons may affect the comparability of rates for these groups. If there is any connection between death of the infant and registration or non-registration of birth, the method will obviously give biassed rates. For example, if the physician on registering a death remembers and records the hitherto unrecorded birth; or if a doctor is called in case of an infant death shortly after birth where in case of survival for the first few days a midwife only is in attendance, the rate may be biassed; for the birth would be more likely to be recorded in case of death than if the infant survived the critical period of the first few days. But if the physician defers registration of birth and thinks it unnecessary if he has to record the death in a few days after birth, the rates will be biassed in the other direction. If these tendencies affect the well-to-do and poor differently or in different degrees, comparative rates may not tell the true degree of difference in rates between these classes. The method purports to be an unbiassed method of selection; it is unbiassed as far as personal or arbitrary bias of the investigator is concerned; but it may actually be biassed in favor of low rates for the entire group or for certain subgroups. It is therefore necessary even with this method to test the presence and determine the character of bias.

Several of the studies of the Bureau have been made on the basis of registered births: Manchester, Brockton, and New Bedford. An illustration of the effect of bias in the method of selection is afforded by Saginaw; the mortality rate for infants of foreign born mothers in the lowest income group came out approximately 50 per 1,000, nearly as low as the rate for infants of native mothers of the highest income group. When the infants that had been excluded on account of non-registration of birth were added the rate was much higher and more in accordance with what other studies have shown. Another peculiarity of the Saginaw material is the fact that

in the group of infants of foreign born mothers there were no deaths of infants aged less than 1 day reported. The usual percentage is about 15, and though the total number is small (only 20) the absence of deaths under one day is a trifle suspicious. That no stillbirths were reported by midwives may possibly be due to erroneous or neglected registration.

The third method assumes in absence of evidence that both birth and death records are likely to be deficient and attempts by the best methods available to remedy the deficiencies. Whether all births and deaths would be discovered would depend on the scope and carefulness of the canvass.

In a word, the first method usually errs in overstating the mortality rate; the second and third methods attempt to secure more accurate rates, the second by selection of births and the third by supplementing the data.

The choice of method is frequently limited by the practise of the registration office. In Waterbury and Baltimore, for example, if an infant death was recorded where the birth had occurred in the city, the birth was immediately recorded if not previously registered. The distinction here, between the first and second methods, became impossible to make. In several cities, therefore, and especially in those cities with faulty birth registration, the records have been supplemented by a canvass or in other ways. In Saginaw, 149 live births were added to total of 1,113 even without a systematic canvass, by use of baptismal records and neighborhood inquiry. In Waterbury, 329 unregistered live births were added to the 2,239 registered. In Akron (outside of the birth registration area), a canvass was obviously necessary. In Baltimore, a campaign for complete birth registration by all the civic and social agencies was being conducted during the progress of the infant mortality study with the result that many births were added.* The tendency, therefore, has been to resort to the method of supplementing official records by a canvass.

In this connection the exclusion or separate treatment of migratory families is the more justified because it fits in with the plan of securing additional information by a canvass and

* In Brockton and New Bedford a canvass is made by special state agents to correct registrations; additional births discovered are added to the registered births.

simplifies the problem very greatly. If the group to be studied includes only cases where the infant is born in the year selected and the family is living in the city at the time of the canvass it becomes at least possible with a thorough canvass to locate all births and deaths that ought to be included. The death records need be supplemented only by deaths of infants born to mothers who are still living in the city, normally these deaths will be recorded. The birth records are to be supplemented by the addition of any children living in the city at the time of the canvass, who were born in the city in the selected year. Of course some difficulties still exist. But in the main the procedure is simple and the data for this "canvassable" group are reasonably complete and accurate.

In any of these methods there arise interesting and often perplexing problems of identification. Sometimes an infant's name is spelled in one way on the birth certificate and in another on the certificate of death.* The phonetic spelling of the name by the doctor who makes the return may not agree at all with the spelling preferred by the family of the child. These cases are fairly common among the foreign groups, especially with the perplexing names of Poles and Lithuanians. It is sometimes difficult to locate families even when they have not moved, as in one section of Waterbury where the foreigners —this is a free country indeed—chose lucky numbers for their houses. Families that move once or twice to other places within the city are often hard to locate because neighbors do not always remember the new address. If a canvass of the city is made these families are found at the correct addresses by the canvassers.†

Turning to the analysis of results, in the Baltimore material we have for the first time in these studies a large enough number of colored infants to give a significant mortality rate. The rate for white infants is 95.9; for colored, 158.6, or 65 per cent higher.‡ The foreign born in Baltimore have the same rate as the native; but analysis of the foreign born group show

* For example, Zartsch and Sage, Greenwood and Boisvert.

† There are always a few cases that can not be traced, for whom it can not be determined even that they have moved away. These cases are omitted, the chances being that they belong rather to the migrant than to the resident group, especially if a canvass is made and fails to locate them within the city.

‡ All figures for Baltimore are provisional.

striking differences among the different nationalities, differences that have been revealed in other studies. The Hebrews have the remarkably low rate of 51.0, the Poles, on the other hand, the high rate of 163.2.

For fathers' earnings groups the rates decrease from the rate (maximum) of 156.7 for the group with earnings under $450 to 37.2 (minimum) for infants whose fathers earned $1,850 or over.

TABLE I.
INFANT MORTALITY RATE BY FATHERS' EARNINGS.

Earnings of Father	Infant Mortality Rate	Live Births	Deaths
All classes	103.5	10,797	1,117
Under $450	156.7	1,544	242
$450–$549	118.0	1,449	171
$550–$649	108.8	1,489	162
$650–$849	96.0	2,417	232
$850–$1,049	71.5	1,595	114
$1,050–$1,249	66.6	661	44
$1,250–$1,449	74.0	419	31
$1,450–$1,849	86.3	371	32
$1,850 and more	37.2	431	16
No earnings	207.7	207	43
Not reported	140.2	214	30

The effect of the kind of feeding can be discussed intelligently only in connection with the age of the infant in intervals of less than a year because of the changes in the proportion of infants artificially fed at different ages. In the following chart for Baltimore the monthly death rates are shown for artificially fed and breast fed infants and the probability of dying in the month is given for all infants except those not fed. The rates for the artificially fed are from three to nine times those for breast fed infants. The relative percentage disparity between the rates is greatest from the third to the eighth months and is greater for these than for the first month. In this respect the figures agree with the data for Berlin for 1895–96, though there the rates for the artificially fed were considerably higher than in Baltimore.

Besides the two main groups there is a group of infants with mixed feeding. The rates for this group lie midway between the rates for the other groups; for the first month it is nearly equal to the rate for the artificially fed, but for later months it approximates more closely the rate for breast fed infants. For all of these groups the type of feeding taken is that for the major part of the month, or, which amounts to the same thing,

TABLE II.

MONTHLY PROBABILITY OF DYING PER 1,000 SURVIVING AT BEGINNING OF SPECIFIED MONTH OF LIFE AND MONTHLY DEATH RATES BY TYPE OF FEEDING IN BALTIMORE, 1915.

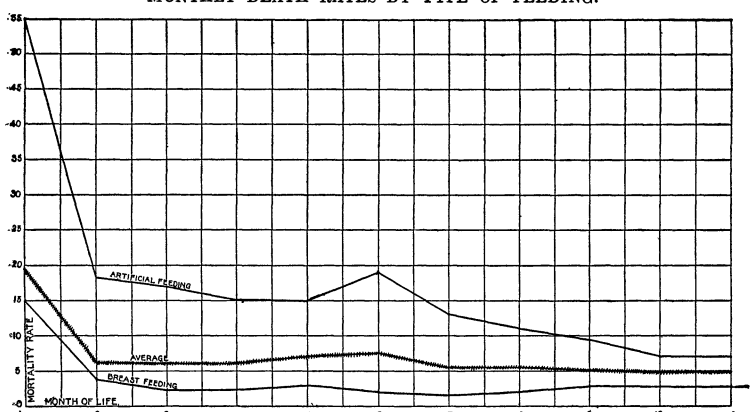
Month of Life	All Infants	Breast Fed	Mixed	Artificially Fed	Relative Difference*
1.....	(44.1) 19.8†	15.0	42.7	55.3	3.7
2.....	6.3	3.8	4.9	18.9	5.0
3.....	6.1	2.4	9.5	17.0	7.0
4.....	6.1	2.3	5.2	15.3	6.6
5.....	7.0	3.0	5.6	15.1	5.0
6.....	7.6	2.1	4.0	19.1	9.1
7.....	5.6	1.7	3.2	13.0	7.7
8.....	5.6	2.2	3.3	11.2	5.1
9.....	5.2	2.8	3.1	9.5	3.4
10–12.....	14.3‡	8.2‡	7.0‡	28.8‡	3.5

* Relative difference between artificial and breast feeding, column 4 divided by column 2.

† Figure is for probability of dying for infants who lived long enough to be fed, the figure in parentheses is the probability of dying for all live born infants.

‡ Figures are, first column, probability of dying before end of year per 1,000 infants surviving at beginning of the tenth month; second to fourth columns, deaths in tenth to twelfth months per 1,000 infants at risk at beginning of period

CHART I. MONTHLY PROBABILITY OF DYING IN THE FIRST YEAR OF LIFE, AND MONTHLY DEATH RATES BY TYPE OF FEEDING.*



* Figures in Table II.

for the middle day of the month. The number of infants that died breast fed is therefore divided by the approximate number of months lived by breast fed infants for the month of life considered, giving a monthly death rate for this type of feeding.*

* The information for infants for whom the type of feeding changed during the month is scarcely so accurate as suggested in the discussion. The above treatment would give the correct result if the data were gathered in every case according to instructions. In cases where instructions were not followed, the above treatment would be as safe as any. Such cases would form a relatively small proportion of the total.

TABLE III.

DEATHS PER 10,000 LIVING AT EACH AGE ACCORDING TO TYPE OF FEEDING: BERLIN, 1895–96.*

Age in Months	Breast Milk	Cows' (and Goats') Milk	Column 2 Divided by Column 1
0	201	1120	5.6
1	74	588	7.9
2	46	497	10.8
3	37	465	12.6
4	26	370	14.2
5	26	311	12.0
6	26	277	10.7
7	24	241	10.0
8	30	213	7.1
9	30	191	6.4
10	31	168	5.4
11	39	147	3.8

* Except last column, from Westergaard, H., *Die Lehre von der Mortalität und Morbilität*, 2nd Ed., Jena, 1901, p. 361. Results of Böckh for Berlin, based on census of infants in city (.2 unknown) and death certificates—(exclusive of infants who died in first day, the type of feeding was unknown in one seventh of deaths, for 1896, these were distributed pro rata of known cases.)

The results of the table for Baltimore may be stated in other words. If the rate for the artificially fed had been the same as that for the breast fed infants, only 92 instead of 437 deaths would have occurred among the former group. The rate for breast fed was only 21.3 per cent. of that for the artificially fed on an average. If all the infants had been breast fed and the same rate had applied to all as characterized the breast fed infants the total deaths of the group studied would have been reduced by 345, from 1,117 to 772, the rate from 103.5 to 71.5 per 1,000 live births.

Of course some of the artificially fed babies may be weaker than their more fortunate brothers. One might assume the effect of such weakness, if it necessitated artificial feeding, would be especially marked in the first month of life. But the relative difference in the rates is greater for the third and later months than for the first and second. Light may be thrown on the effect of continued artificial feeding and of a shift of feeding by an analysis of mortality rates by duration of each type of feeding.

Analysis of mortality rates for each type of feeding by economic conditions might reveal clearly the effects of other elements in mortality. As far as nativity of mother is concerned the evidence shows in most cases that a larger proportion of native mothers than of foreign born mothers feed their infants artificially. The differences between the average rates for the

two groups, therefore, is somewhat minimized by neglecting to make a secondary analysis by nativity classes. A difficulty in interpreting the results of such a secondary analysis, however, is that artificial feeding does not always have the same implications. The native mother may be more careful in preparation of the cows' milk for the baby, she may make use of the formulae for modifying the milk that insure the best results. This difference in the meaning of artificial feeding for the nativity groups must be taken into consideration in interpreting the conclusions.

Another problem is the question of the effect upon mortality of employment of the mother in the year following the birth of the infant. It has been shown that a relatively small percentage of mothers are gainfully employed away from home. In Baltimore, mothers of 859 infants, or 8.0 per cent. of the total, were employed away from home during the lifetime of the infant. In Waterbury only 3.4 per cent. of the infants had mothers so employed, an even smaller proportion of the total. The general mortality rate for the city could not be primarily influenced or much affected by a high mortality among infants whose mothers were gainfully employed. It is quite possible, however, that gainful employment of the mother away from home may increase considerably the mortality rate *among the infants of these mothers.* To determine this effect it is necessary to know the age of the infant at the time the mother went to work.*

In Baltimore, of these 859 infants whose mothers were gainfully employed away from home, 78 died. If the rate for all infants had been applied to these infants at the ages at which the mothers went to work only 39 deaths would have resulted. In other words the mortality rate of infants whose mothers were gainfully employed during the lifetime of the infant was about twice the average rate.

Further analysis might determine whether this excess mortality was due merely to the necessity of artificial feeding or to the preponderance among the group of infants of colored mothers or of infants in the low income groups. If gainful

* If the infant died before the mother went to work or resumed employment, the death can not be attributed to the employment of the mother.

employment of the mother involves artificial feeding the mortality among the infants would normally be high; artificial feeding, however, may not be sufficient to account for the difference. If a large majority of those gainfully employed were colored, one would expect more deaths than if the mothers were white; and similarly if a large proportion were in the relatively poorer classes, as undoubtedly was the case. Extended analysis may indeed show, as suggested by Hibbs,* that the earnings of the mother, by providing means for extra care and somewhat better conditions, may more than make up for the handicap under which the infant suffers through absence of the mother during work hours. By calculating the deaths that would have occurred among the infants of mothers gainfully employed for each such group at the rate for the subgroup (colored, white, poor, well-to-do, etc.), applied to the corresponding numbers in it, first with the rate for breast fed infants and then with that for artificially fed infants, it would be possible to show how much handicapped these infants are by reason of not being breast fed and how far, if at all, they are favored over the artificially fed infants of mothers who are not gainfully employed.

These studies throw light upon the connection between the death of the mother and infant mortality. It is to be expected that when the infant is deprived of the mother's care the mortality would be higher. In Waterbury there were 16 live births to mothers who died during the year following confinement. At the average rate of mortality for the city (122), two deaths would have occurred among them but actually six deaths occurred, or three times as many. In Baltimore there were 68 live births to mothers who did not survive the next year. At the average rate for the city 7 deaths would have occurred; there were actually 34 deaths or nearly five times as many.†

A very interesting fact is brought to light in examining the proportion of families of a given number of births that had no

* But not demonstrated; the figures quoted are for the Johnstown report of the Children's Bureau, in which no analysis is made of the age of infants at the time of the mother's resumption of employment. The figures then merely compare infant mortality in families where the mother works with that in families where the mother is not employed; this comparison is an essentially different proposition and is not so directly related to the problem. Hibbs, Henry H., Infant Mortality: Its Relation to Social and Industrial Conditions, p. 108.

† Cf. Westergaard, H., *Die Lehre von der Mortalität und Morbilität*, 2nd. Ed., Jena, 1901, p. 376.

infant deaths. In families of two births if the mortality rate is 100 or one in ten die, the probability that both children die before the end of the first year is one in 100, and that both children survive the first year 81/100. Similarly for families of n births, if q is the probability of surviving the first year, q^n is the proportion of families where all infants survive. A comparison of the actual and expected rates as so calculated reveals the interesting fact that there are more families with no deaths than would be expected by the laws of chance: in other words the deaths are found to be slightly concentrated in relatively few families. That this result is not strange follows directly from the consideration that when the care given by the mother is unusually good, or when the physique inherited by the infant safeguards him through his first year of life, it is likely that his later brothers and sisters will enjoy the same advantages as he. The occurrence of infant deaths among the later births is not entirely independent of the causes leading to deaths among the earlier births in the same families.

A study of the effect on mortality of the interval between births is permitted, though not as directly as might be wished, by a classification of births by age of mother and order of birth. Where the mother is under 20, and two or more births have occurred, it is highly probable that the interval between births has been short; by a combination of age of mother and order of pregnancy, it appears that the rates for the later numbers in order of pregnancy advance rapidly, and always more rapidly the younger the age of the mother in the group considered. For third births, for example, the rate is 196 for mothers under 20, 129 for mothers 20–24, 99 for mothers 25–29, etc. For fifth births it is 184 for mothers 20–24, 134 for mothers 25–29, and 95 for mothers 30–34. Of course it is possible that a subclassification by color or economic conditions might explain some of those differences.

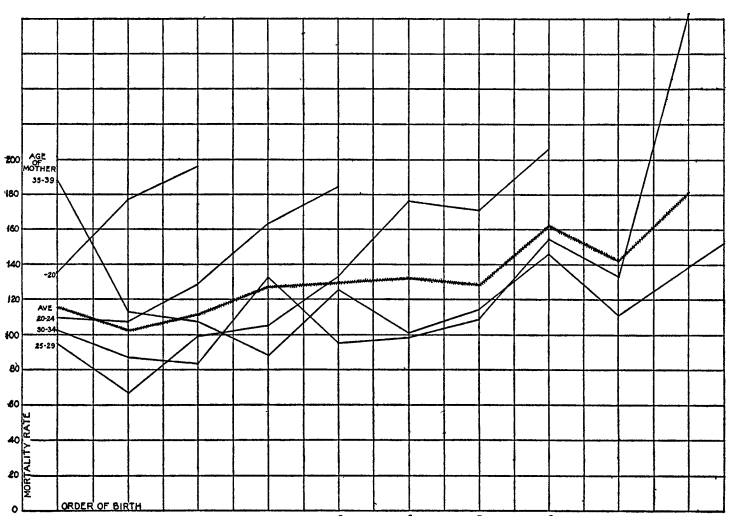
TABLE IV.

MORTALITY RATES BY AGE OF MOTHER AND ORDER OF BIRTH: BALTIMORE.*

Number in Order of Birth	Average	Age of Mother					
		Under 20	20–24	25–29	30–34	35–39	40 and Over
Average		148.1	118.6	107.7	112.8	126.7	131.2
1.....	115.8	135.1	109.8	95.1	102.4	(188)	.....
2.....	102.7	177.2	107.5	66.8	87.0	113.2	.....
3.....	111.5	196.4	128.7	99.4	83.5	107.6	.....
4.....	127.0	(290)	163.4	105.1	132.4	88.4	.....
5.....	129.3	.....	184.0	133.6	95.4	125.5	.....
6.....	132.2	.....	(179)	176.7	98.2	101.1	.....
7.....	128.2	.....	.....	171.6	108.7	114.5	.....
8.....	162.6	.....	.....	203.7	154.5	146.1	.....
9.....	142.1	.....	.....	(295)	133.0	111.1	.....
10.....	181.1	.....	.....	.....	282.3	138.4	.....
11.....	146.8	.....	.....	.....	.....	152.4	.....

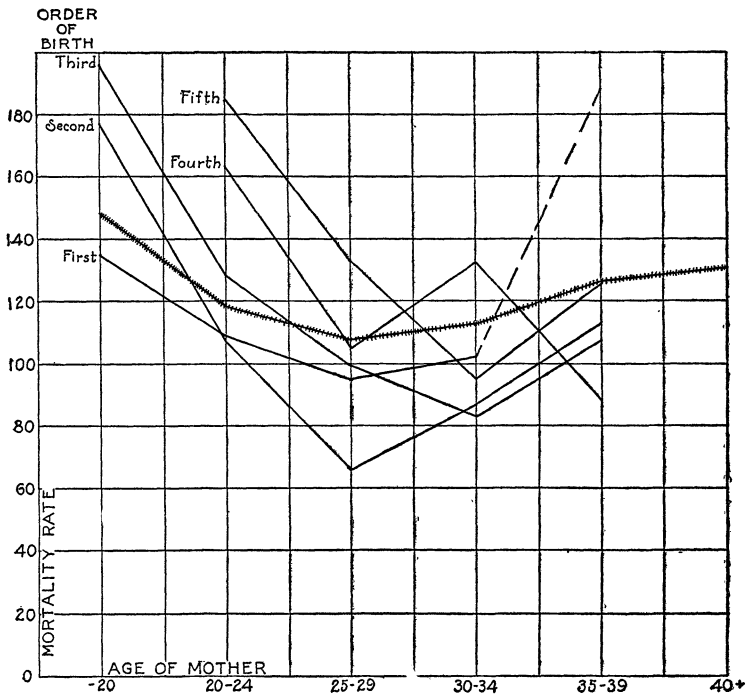
* Includes all births reported by mothers of infants born in Baltimore in 1915.
Rates bracketed based on fewer than 100 cases. Rates for twelfth and later births omitted.

CHART II. GRAPHS FOR EACH AGE-OF-MOTHER GROUP SHOWING VARIATION IN INFANT MORTALITY RATES BY ORDER OF BIRTH.*



* Figures in Table IV.

CHART III. GRAPHS FOR EACH ORDER OF BIRTH SHOWING VARIATION IN INFANT MORTALITY RATES BY AGE OF MOTHER.*



*Figures in Table IV.

The effect of lack of proper spacing between births can be studied by analysis of mortality rates for families of different sizes (number of births to mother) and average duration of married life. That the interval between births is short is obvious when the "terms" of married life are short and the number of births is relatively large; but for long durations of married life it is uncertain whether births have been properly spaced or not.

The results for Waterbury are shown in the accompanying table, giving the mortality rates for infants of mothers married less than 6 years, 6 to 10 years, 11 to 15 years and 16 years and over. In the first three groups the tendency is marked for the

rate to increase as the interval, indicated by a large number of births, decreases.

The second table is an attempt to group the results more clearly in terms of the interval between the birth and the beginning of the succeeding pregnancy. The rates where the average interval between the birth of one child and the next pregnancy is approximately 1 year is considerably higher than where the interval is approximately 2 years or 3 years or over, 151.4, 111.2, and 102.3 respectively. An interesting feature of the table is the decided increase in the rate as the number of successive births following one another by intervals of less than 2 years increases. That the result is not explained by the increase in size of family alone is indicated by the fact that in large families where the average interval is over two years there is but little difference in rate from that for the smaller families. One might fairly conclude that an important cause of high mortality is the lack of proper spacing of births.*

TABLE V.

MORTALITY RATES BY LENGTH OF MARRIED LIFE AND NUMBER OF BIRTHS TO MOTHER: WATERBURY.

Number of Births	Length of Married Life											
	5 Years and Under			6–10 Years			11–15 Years			16 Years and Over		
	Live Births	Deaths	Rate	Live Births	Deaths	Rate	Live Births	Deaths	Rate	Live Births	Deaths	Rate
1...	469	47	100.2	11	0	0.0	...	..	.....	...	..	.....
2...	656	61	93.0	153	18	117.6	26	2	(77)	2	0	.....
3...	510	73	143.1	408	38	93.1	54	4	(74)	15	1	(67)
4...	146	28	191.8	676	82	121.2	157	16	101.9	46	6	(130)
5...	15	5	(333)	552	83	150.4	355	32	90.1	71	7	(99)
6...	...	..	.....	255	48	188.2	429	43	100.2	83	5	(60)
7...	...	..	.....	98	20	(204)	398	46	115.6	200	34	170.0
8...	...	..	.....	21	4	(190)	280	47	167.9	280	41	146.4
9...	...	..	.....	9	0	0.0	137	31	226.3	269	32	119.0
10..	...	..	.....	...	..	.....	60	19	(317)	314	49	156.1
11..	...	..	.....	...	..	.....	11	5	(455)	135	24	177.8
12..	...	..	.....	...	..	.....	12	8	(667)	140	24	171.4

Rates for 13 and over omitted.

* Part of the differences in rates may be due to the tendency for the interval between births to be shortened in case of death of an infant. See page 47.

TABLE VI.

MORTALITY RATES BY APPROXIMATE AVERAGE INTERVAL BETWEEN BIRTH AND NEXT PREGNANCY BY NUMBER OF BIRTHS TO MOTHER: WATERBURY.

Number of Births	Average Interval 1 Year				Average Interval 2 Years				Average Interval 3 Years or Over			
	Years Married	No. Live Births	Deaths	Rate	Years Married	No. Live Births	Deaths	Rate	Years Married	No. Live Births	Deaths	Rate
2.....	– 4	590	55	93.2	5–6	141	14	99.3	7+	106	12	113.2
3.....	– 6	660	81	122.6	7–9	229	25	109.2	10+	98	10	(102)
4.....	– 7	562	89	158.4	8–11	307	26	84.7	12+	156	17	109.0
5.....	– 9	486	78	160.5	10–14	411	47	114.4	15+	96	8	(83)
6.....	–11	356	57	160.1	12–17	357	36	100.8	18+	54	3	(56)
7.....	–13	392	54	137.8	14–20	246	25	101.6	21+	28	1	(36)
8.....	–14	237	41	173.0	15–22	325	45	138.5	23+	19	6	(316)
9.....	–16	189	34	179.9	17+	222	29	130.6	....	...	..	.....
10.....	–18	217	49	225.8	19+	157	19	121.0	....	...	..	.....
11.....	–20	74	20	(270)	21+	72	9	(125)	....	...	..	.....
12.....	–21	82	24	(293)	22+	77	8	(104)	....	...	..	.....
2–12....	...	3,845	582	151.4	.....	2,544	283	111.2	....	557	57	102.3

Rates for 13 and over omitted; (6 mothers) 16 deaths, 77 live births.

A third method for the study of the effect of interval between births is available but hitherto not used. The schedules gathered show usually the age of the infant born in the selected year at the commencement of a later pregnancy as well as a rough statement of the interval between the birth of the child in the selected year and the preceding birth. By analysis of these it may be possible to determine the effect on mortality of the child following another by a short interval as well as of the child preceding another by a short interval: the effect in the two cases may be of different character and of different magnitude. By excluding infants that died before the commencement of next pregnancy, this method of procedure would have the obvious merit of eliminating the effect of a tendency, revealed in European statistics, for the interval between successive births to be shortened in case of the death of an infant at an early age.*

There remain two very important problems of treatment of data, both of which have been suggested in the foregoing.

One is the problem of determining how many cases suffice to establish a conclusion. In the example cited, p. 42, is the fact that 34 deaths occurred, instead of 7 expected, among the 68

* Cf. Westergaard, H., *Die Lehre von der Mortalität*, 2nd ed., pp. 370–1.

live births a sufficient basis for a conclusion that the death of the mother exercises an adverse effect on infant mortality? The problem may be stated in slightly different words: is the result due to a significant cause or causes or is it to be explained merely as a chance combination of insignificant or unimportant causes? If we can show that it is not explicable on ground of "chance combination of individually unimportant causes," we have taken a long step toward establishing the true causal connection, though it may or may not be true that the cause alleged is the important factor.* The procedure is simply to calculate the probability or improbability of the result occurring by chance, by analogy with the results of games of chance.

The proposition might be put as follows: If black and white balls are mixed in an urn in the proportion of 103 black to 1,000 total (the mortality rate in Baltimore was 103) and 68 are drawn at random, the most probable number of black balls would be 7. How frequently would 34 or more of the balls be black? It is easily seen and it can be mathematically shown that such a result is extremely improbable—it would occur but a few times in a million drawings. One may conclude, therefore, reverting to the problem under discussion, that some significant causes are acting to increase the mortality of infants in cases where the mother died, that there is a causal connection between the two.

The mathematical formula for the root mean square error (or, if preferred, the probable error) may be used to test the probability of a result occurring from a chance combination of of insignificant individual causes in connection with the differences between the observed and expected results. A deviation of three or four or more times the standard or mean error makes it practically certain that real causes are acting.†

The number of cases necessary to establish a conclusion depends (a) on the probabilities involved, and (b) on the amount of the differences between observed and expected results.

The second problem is that of separating the effect of two or

* Compare the discussion on the effect on mortality of gainful employment of mothers, p. 41.

† The applicability to mortality rates of this procedure of comparison by analogy with results of game of chance has been discussed by Westergaard and others. Cf. Westergaard, H., *Die Lehre von der Mortalität*, 2nd Ed., pp. 186 ff., esp. 197–9. *Die Grundzüge der Theorie der Statistik*, ch. 5. Scope and Method of Statistics, QUARTERLY PUBLICATIONS, American Statistical Association, Vol. XV, pp. 244 ff., Sept., 1916.

more interrelated factors and assigning to each its proper and due weight. The problem may be made more concrete by discussion of some specific cases, for example, age of mother and order of birth, or race and earnings of father.

The heavy lines in the charts show the variation in infant mortality by age of mother and by order of birth separately. But it is possible that the high rates for mothers over 40 is due to the large proportion of births of late orders to such mothers —births, among whom the mortality rate is high, or vice versa. How can these effects be separated?

Various methods have been suggested, the correlation coefficient, multiple correlation, some method of standard distribution, and Westergaard's method of expected and actual deaths.

Analysis of the problem shows that the method to be preferred depends upon the amount of information available. The object sought may be either to sum up in easily understood form the data of a complex table, or it may be to add to the information at hand by further analysis.

Suppose, for example, that infant mortality rates are at hand for various cities and that nothing is known of mortality rates for infants of native and foreign born, of illiterate and literate mothers, of mothers employed and not gainfully employed. Some light may be thrown on the relationship by securing indices of these factors for the different cities and classifying the cities by the size of the indices and average mortality rates, in other words by correlating mortality with the various indices. It will be readily admitted that the conclusions are likely to be precarious, first, because comparatively few cities may be used, second, because indices may not be very directly related to the problem investigated, and thirdly because important factors may be overlooked. A correlation may be found between infant mortality rates and the percentage of foreign born, which may be due solely to the poor registration of births in cities with a large percentage of foreign born. Where complex relationships, such as between the mortality rates and age of mother and order of birth, are investigated by this method, the selection of a true index becomes almost impossible, especially in case the relationship is not linear. Yet in absence of

classified births and deaths, such a method of analysis may be better than nothing.

A second case is where it is sought by further analysis to determine which of two factors is the more important causally and to determine the limits of the effect of each. This problem appears in two forms: where the only available information is a series of rates for each factor separately, and where rates are given for each combination of both factors. In the latter case the problem is one of combining results and presenting the significant conclusion. But in order to determine the method of approach and the possibilities of analysis where only part of the information is available, it will be pertinent to give a few illustrative cases where conclusions based on partial data may be tested by the fuller information.

One difficulty in the interpretation of average rates for each factor separately is that the average may be influenced by the special distribution of births in each group. A second difficulty is the interpretation of causality. In the example under discussion it is probable that neither age of mother nor order of birth are per se "causes," but simply that they connote or imply conditions that favor high mortality—impaired physique, or lack of experience. Low age of mother may obviously connote a low order of birth. To the extent that one connotes the other, one is as significant as the other, except that one may be shown to be more directly related to the true cause. All that statistical analysis can show at best is the extent to which each factor connotes something not connoted by the other, and the extent to which both factors have the same connotations.

CASE 1. FACTORS INSEPARABLE.

(1,000 births assumed in each group where rate is shown.)

Age of Mother	Order of Birth.					
	Average	1	2	3	4	5
Average		100	110	120	130	140
20	100	100				
22	110		110			
25	120			120		
32	130				130	
38	140					140

CASE 2. AGE OF MOTHER SOLE CAUSAL FACTOR.

Age of Mother	Order of Birth					
	Average	1	2	3	4	5
Average		105	110	120	130	135
20	100	100				
22	110	110	110	110		
25	120		120	120	120	
32	130			130	130	130
38	140				140	140

CASE 3. ORDER OF BIRTH SOLE CAUSAL FACTOR

Age of Mother	Order of Birth					
	Average	1	2	3	4	5
Average		100	110	120	130	140
20	105	100	110			
22	110	100	110	120		
25	120		110	120	130	
32	130			120	130	140
38	135				130	140

CASE 4. BOTH FACTORS INFLUENTIAL.

Age of Mother	Order of Birth					
	Average	1	2	3	4	5
Average		105	120	140	160	175
20	105	100	110			
22	120	110	120	130		
25	140		130	140	150	
32	160			150	160	170
38	175				170	180

The average rates for the different orders of birth and for the different ages of the mothers show the same general trend in all four cases. It seems clear from a study of these examples that where merely the averages are given it is impossible to determine the relative effect of each factor.*

* Size of correlation coefficient appears to be inconclusive, if only averages for each factor are available.

If in addition to the average rates the distribution of births classified by both factors is given, it may be possible to throw some light upon the relative influence of each cause.* Perhaps the most satisfactory method is that suggested by Westergaard of calculating deaths on the assumption first of a uniform rate for each order of birth, and then for each age of mother, comparing the calculated with the actual deaths to determine the trend of the true differences. The method has its limitations. But applied to the cases given above, it indicates correctly (except for the first, which is indeterminate) which is the true causal factor, though perhaps minimizing its effect. The method is applicable even if the relationship between the rate and either or both factors is not linear. If another cause, such as interval between births, is more important than one of the factors under discussion the result is of course inconclusive to the extent that this cause has been disregarded.

If in addition to births the deaths are also subclassified, it becomes possible to give rates for each subgroup and to determine definitely the causal influence of each factor. The problem here is to combine the results in some easily understood form that epitomizes the data, showing the true weight of the causes considered. But for the statistician nothing can quite replace the rates themselves. For if some other cause is of considerable importance besides those presented, a summary process may obscure its effect and lead to erroneous conclusions. In the case under examination it becomes evident from the study of the charts showing the variation of rates for each order of birth by age of mother that the interval between births affects the mortality rates by age.

If the two factors are the only important ones to be considered, and the object is merely to present a result not affected or warped by a peculiar distribution of births, the best method is probably some form of the method of standards. For example, the difference between the rates for colored and white is somewhat exaggerated by the large proportion of colored in the lowest economic group. It is possible to present the true

*The character of the distribution of births often permits of approximately correct conclusions from the averages of two factors; it is where the distribution is "warped" that the conclusions need to be corrected. The method of a standard distribution is merely a substitution of a symmetrical for a "warped" distribution.

difference between colored and white either by calculating deaths in a *standard population* by applying first white and then colored rates: the sum of the deaths divided by the population gives the standardized rates. But it is rather better to use the actual distribution of population in the group as the *standard itself;* and calculate the deaths that would have occurred among the colored population if the rates prevalent among the white population had applied. The advantage of this method is simply that undue weight is not given to those rates which are based on relatively few cases, and that it is possible to present the number of cases upon which the rates are based so that the standard error may be calculated.